

**RULE INDUCTION FOR SUMMARIZING
THE CLASSES IN A CLASSIFIED
DOCUMENT COLLECTION**

ABSTRACT OF THE DISCLOSURE

5 A method and apparatus for providing summaries
of documents belonging to a class of documents in a
classified document collection. A sample set of
documents belonging to one or more classes is
processed via a machine learning system in order to
10 induce a set of rules associated with the sample set
of documents. The vocabulary in the rules are
extracted and compared to words, terms or phrases of
an incoming document. Any matches between the
extracted rules and the words, terms or phrases of
15 the incoming document are used as a summary for the
incoming document. By using the method and
apparatus, each document does not have to be
processed to find most important words and the like
in order to provide a summary for that document and
20 then repeating the same process for additional
documents.